U.S. Department of Transportation
United States Coast Guard

Commandant
United States Coast Guard

2100 Second St. S.W.
Washington, DC 20593-0001
Staff Symbol: G-SIA
Phone: (202) 267-1798

COMDTINST 5230.42A
MAY 17 1996

COMMANDANT INSTRUCTION 5230.42A

Subj: COAST GUARD DATA ELEMENT NAMING STANDARDS

Ref: (a) Information Resource Management COMDTINST 5230.41
(b) Planning Approval For Automated Information Systems COMDTINST 5231.2

1. PURPOSE. To promulgate procedures and assign responsibilities for compliance with Coast Guard Data Element Naming Standards and the enforcement of these standards.

2. ACTION. Area and district commanders, commanders of maintenance and logistics commands, commanding officers of Headquarters units, Headquarters Chiefs, Directors; and special staff divisions involved in an information systems project shall comply with the contents of this instruction.

3. DIRECTIVES AFFECTED. Commandant Instruction 5230.42 is superseded.

4. BACKGROUND.

a. The Coast Guard recognizes data as a corporate resource. The management of that data is critical to the success of Coast Guard missions. A coherent set of naming standards providing for the uniform identification, definition and representation of data elements is crucial to the central management of Coast Guard data.

b. The Coast Guard has an ever increasing need to share data across organizational and functional boundaries. A Cross-Functional System (CFS) is an information system that supports organizational processes relating the activities of several programs or functional divisions, rather than activities of a single program. Development of these CFSs add further emphasis and dimension to the need for responsive standards facilitating this data.

5. DISCUSSION. Data element naming standards support the sharing of Coast Guard data resources through uniform data representation, consistent interpretation and common understanding throughout the organization. Enclosure (1) is the Coast Guard's Data Element Naming Standards (DENS) Guidebook. It lists and describes the rules for structuring standard data element names and definitions (these rules are in conformance with both National Institute of Standards and Technology (NIST) and Department of Defense (DOD) data element naming standards). In addition, it explains in detail the importance, methodology and conceptual foundations for establishing naming standards. Specific objectives of Coast Guard data element standards activity are to:

   a. Ensure that data is treated as a Coast Guard-wide shared, controlled and managed resource.

   b. Reduce the cost of managing data by eliminating redundancies.

   c. Provide a single approach to data element standardization for use throughout the Coast Guard. Data standardization procedures provide the documentation needed to coordinate data sharing across the Coast Guard.

   d. Ensure the accuracy and integrity of data throughout Coast Guard information systems.

   e. Improve the quality of future Coast Guard data elements.

   f. Make standard elements available and accessible for satisfying information system requirements Coast Guard-wide sharing.

6. POLICY. The scope of Coast Guard Data Element Naming Standards encompasses data supporting Coast Guard missions and goals, with the exception of stand-alone Coast Guard Automated Information Systems (AIS) that neither share data nor have potential for sharing data with other AIS. An example of this would be a small local unit data base on a standard workstation used strictly for processing data within that unit. This policy and guidance applies to the development of new AIS and those undergoing major redesign/modifications. Offices sponsoring these initiatives

shall comply with the data element definition and naming rules listed in enclosure (1), the Data Element Naming Standards Guidebook. The goal of the Data Administration Dictionary System (DADS) implementation is to standardize data elements used in CFS to support data sharing. Commandant (G-SIA), as the Coast Guard Data Administrator, will manage the data element standardization and approval processes.

a. The system's sponsor (the project manager from the Coast Guard office responsible for the system) and the system's developer shall:

(1) Identify data elements required to support applications as early as possible in the life-cycle management phases of the system.

(2) Submit to Commandant (G-SIA) a list of candidate prime words (see enclosure (1)) for possible inclusion in the DADS.

(3) Construct the data element definitions and standard names in accordance with the rules in enclosure (1).

(4) In addition to the data element name and definition, prepare the following attribute information for each data element:

(a) Data element data type (CHARACTER STRING, INTEGER, FIXED POINT, FLOATING POINT, OR BIT STRING).

(b) Data element maximum length. This is the maximum number of characters that the data element is allowed to contain.

(c) Data element domain. This is the range of values that the data element can have.

(d) Responsible office of the new application.

(e) Name of the new system and/or application.

(f) Alias name(s) of the data element to be used in the new application. A data element's standard name is its logical representation, including descriptive information, across all applications. An alias is the representation, usually significantly abbreviated, of a data element within a particular application.

(5) Enter data prepared in paragraphs (3) and (4) above into the DADS.

(6) Ensure that a contract for the development of a cross-functional (see paragraph (4) above) AIS data base provides for the resources required to perform

instruction. The following is an example of a development contract to request additional services:

EXAMPLE: "The contractor shall adhere to the Coast Guard Data Element Naming Standards and use DADS to collect or retrieve data element information relating to systems which the contractor develops and/or maintains. The contractor shall participate in the identification of standard prime words relevant to the systems which is developed and/or maintained. The contractor shall be required to develop and upon Coast Guard review and approval, institute procedures for use by project teams for submitting and entering information to the repository."

(7) Contact the Coast Guard Data Administrator, Commandant (G-SIA), for guidance on submitting the data element names and standard attributes for approval.

b. Commandant (G-SIA) will:

(1) Provide assistance to project sponsors and developers in understanding and complying with the requirements of the DENS.

(2) Provide guidance on the submission of data element information for approval by the Data Administrator.

(3) Create and maintain a database (the Coast Guard DADS) of all standard data elements and attributes submitted and entered into the dictionary according to paragraph 4.a above.

(4) Review all prime word lists submitted by functional proponents and enter those approved into the database (see previous paragraph).

(5) Make information in the DADS database available to system sponsors and developers. Provide procedures for access to these databases.

(6) Convene the IRM Peer Group, consisting of senior IRM representatives from program offices to review, as necessary, Data Administration (DA) issues. The IRM Peer Group shall recommend a resolution to conflicts concerning the assignment of a responsible office for a data element that crosses functional boundaries; only IRM Peer Group members from offices that may potentially be assigned as the responsible office will participate. The IRM Peer Group shall also recommend resolution to definition and format discrepancies for shared Coast Guard data elements.

also recommend resolution to definition and format discrepancies for shared Coast Guard data elements.

(7) Review and approve data element names which meet the DENS guideline within thirty days of date of submission. If data elements are not approved, notify proponents of the reasons for the rejection and suggest modifications, if any, to the original submission to qualify the data element for approval. Appendix A to enclosure (1) describes the data element standardization life cycle.

(8) Actively support the developers of Coast Guard AIS as follows:

(a) Assist the developers in using the DADS database to research and compare standard data element name entries.

(b) Provide a standard list of Class Words. This list is included as an appendix in the DENS Guidebook.

(c) Assist in installing new standard data elements and element changes.

(d) Make new data representations available to users.

/s/ D. E. CIANCAGLINI
Director of Information & Technology

Encl: (1) Coast Guard Data Element Naming Standards Guidebook

Enclosure (1) to COMDTINST 5230.42A

UNITED STATES COAST GUARD

DATA ELEMENT

NAMING STANDARDS

GUIDEBOOK

COAST GUARD DATA ELEMENT NAMING STANDARDS
TABLE OF CONTENTS

INTRODUCTION 1

1. WHY DATA NAMING STANDARDS ARE IMPORTANT 2
   1.1 THE UTILITY OF NAMING STANDARDS 3
   1.2 WHAT IS IN A NAMING STANDARD? 3
   1.3 BENEFITS 3

2. DATA ELEMENT NAMING METHODOLOGY 4
   2.1 CLASS WORDS 4
      2.1.1 DEFINITION 4
      2.1.2 CLASS WORD APPROVED LIST..... 5
      2.1.3 TEST FOR CLASS WORDS 5
   2.2 PRIME WORDS 5
      2.2.1 DEFINITION OF PRIME WORDS 5
      2.2.2 PRIME WORD EXAMPLES 5
      2.2.3 TESTS FOR PRIME WORDS 6
   2.3 MODIFIERS AND QUALIFIERS 6
      2.3.1 DEFINITIONS 6
      2.3.2 MODIFIER EXAMPLES 7
      2.3.3 TESTS FOR MODIFIERS 7

   2.4 CONSTRUCTION/RECONSTRUCTION OF
      DATA ELEMENT NAMES 8

   2.5 ENFORCEABILITY 8

3. DATA ELEMENT STANDARDIZATION RULES 8
3.1 DATA ELEMENT DEFINITION FORMULATION RULES 8
3.2 DATA ELEMENT NAMING RULES 10

APPENDIX A DATA ELEMENT STANDARDIZATION LIFE CYCLE..... 11
APPENDIX B APPROVED CLASS WORDS LIST..... 12
APPENDIX C ISSUES AND EXAMPLES CONCERNING THE NAMING RULES 15
APPENDIX D GLOSSARY 24

UNITED STATES COAST GUARD
DATA ELEMENT NAMING STANDARDS

INTRODUCTION: This document discusses the development of data element naming standards for use in a Coast Guard-wide data dictionary.

The National Institute of Standards and Technology (NIST) identified opportunities for potential cost reduction and avoidance and recommended implementation of new data standards. NIST work in this area is specifically identified as FIPS PUB 156, the "Information Resources Dictionary System" (IRDS). The Data Administration Branch (G-SIA) has implemented an IRDS-compliant Data Administration Dictionary System (DADS) using the naming standards described in this document. The basic structure of the CG DADS defines data elements as a subset of data entities.

Data entity and data entity name refer to a class of objects (concepts, persons, events or things) of concern to an organization about which information is maintained, usually on a computerized system. Usually there are many more data elements than other types of data entities about which information is maintained in an organization's data dictionary. The CG DADS stores data, or information, about CG data elements and other data entities related to the data elements, including systems, databases, files, reports and data records that use the data elements. The entities themselves are not stored in the dictionary (e.g., it will not contain the actual databases or values of data elements), but information about these entities (e.g., data element **names**, file and database **names**). In data dictionary terminology, this information about data is called "**metadata**".

Data entities are the crucial component of efficient data sharing throughout an organization. The Coast Guard has the capability to share data electronically on an organization-wide basis. Currently, the Coast Guard is in the planning and development stages of several very large automation projects requiring data sharing across functional and/or organizational boundaries. This trend will continue as evolving technology makes software applications and data more portable across various platforms.

Data Element and Data Element Name are subsets of data entities. A data element is defined here as a basic unit of information built on standard structures having a unique

meaning and distinct units or values. Data elements represent the smallest logical division of information that can be identified. The DADS database contains a data element's logical standard name and alias name(s) (defined in the next paragraph). The DADS also contains additional attributes of the data element, such as the names of the systems and databases that use that data element. The names of data elements tend to be more diverse and require more analysis than the names of other data entities. This Guidebook focuses on data **element** naming standards although the same methodology can be used for naming all data entities. Commandant (G-SIA) will establish naming standards for other entities (e.g. system, database, database table/data record names) in the future.

These data element naming standards create a unique **corporate, or primary, name** for each data element. This name is also called the **standard logical name** of the data element. Because the Coast Guard stores information using more than one retrieval method, it may seem hopeless to reconcile all the differing name constraints (most often size restraints) of these diverse applications. This problem is resolved by using one corporate name combined with several **alias names**. Physical aliases are the names required by different software packages and programming languages using the data element. Each application development language has its own set of naming constraints. Over time, standards will evolve for these physical alias names based on the application generator or software language used to develop each application. An alias can also be another logical representation of a data element in a particular application or system (e.g., the logical name used in an application's data dictionary). All alias names shall reside in the DADS for use in cross-referencing the corporate names they represent. The naming standard rules addressed in this document apply only to the corporate name of the data element, i.e., the representation of the data element in the CG DADS.

SECTION 1: WHY DATA NAMING STANDARDS ARE IMPORTANT

A coherent set of standards for naming data elements is vital to improving the management of data. With the development of the data administration function within the organization, centralized oversight of data becomes possible. Among the advantages of this activity is the reduction of redundant data through consolidation of synonymous and overlapping data elements. This can be achieved by the application of rules which lead to the creation of consistent names.

1.1 **The Utility of Naming Standards**:

The mission of Data Administration is the management of data. **Data is a corporate resource**. While the majority of all corporate data should be available to appropriate users, a centralized Coast Guard command is responsible for the coordination and control. The Systems Directorate Architecture and Planning Office (G-SIA), in its data administration function, provides the strategy and tools to accomplish goals such as the cost reductions for data collection, storage and usage, and the reduction of errors in the processing and design of data.

We cannot manage data that we cannot identify. We cannot identify data that we cannot name. A strategy adopted by G-SIA is the composition of a set of data element naming standards. Used in conjunction with a logical database design, naming standards can provide greater efficiencies in data handling and less confusion as to the specific use and meaning of a data element.

1.2 **What is in a Naming Standard?**

There are two areas of concern when naming data: content and format of the data element name. Content relates to the **essential meaning** of the words chosen for the data name; format relates to the **structure** of the data element name (i.e., its size, shape and the arrangement of the words in the data name).

1.3 **Benefits**:

Naming of data elements is an initial step in developing a data element dictionary. Opportunities identified for cost reduction and avoidance related to data naming standards are:

- Improve identification of existing, valuable information resources that can be used more efficiently by others in the organization, or shared with other organizations
- Simplify software and data conversion through the provision of consistent documentation
- Increase portability of acquired skills resulting in reduced personnel training costs
- Facilitate development, modification, and maintenance of manual/automated systems throughout their life cycles
- Support an organizationally-defined data element standardization program and reduce the unnecessary

development of computer applications when suitable applications already exist

- Support records, reports and forms management, spanning the range from non-automated to fully automated environments

SECTION 2: DATA ELEMENT NAMING METHODOLOGY

The naming standards used in this document adhere to the latest NIST direction.

The basic formula for **logical data element names** calls for:

- **One Prime Word (PW)** describing the object being characterized.
- **One Class Word (CW)** describing the use or purpose of the data element.
- **One or more Modifiers (M)** of the Prime Word or Class Word to make the data name more meaningful.
- **One or two Optional Class Word Qualifiers (Q)** to further categorize the Class Word.

Adhering to the following format creates a unique and descriptive representation for each data element in the DADS database:

DATA ELEMENT NAME = M : PW : M : CW : Q

Prime Words and Class Words will be defined and approved for entry into the CG DADS. Their source will be centrally controlled to establish a restricted vocabulary for consistent naming. They will be used in constructing new data element names and for keyword searches of existing elements.

The Prime Word must be positioned as either the first or second word in the data element name. The Prime Word must also be positioned in front of the Class Word within the data element name.

Each of the basic components (Prime Words, Class Words, Modifiers and Qualifiers) are defined and explained in the following sections. Tests and examples are provided.

2.1 Class Words

2.1.1 Definition: A Class Word is the most important **noun** in a data element name. Class Words identify or describe

the use or purpose of a data element. Class Words designate the type of information maintained about the object (Prime Word) of the data element name.

2.1.2 Class Word Approved List: Although Class Words are common among all organizations and are typically limited to a few words, they should be defined to suit the particular needs of the CG. In constructing CG data element names, an official Class Word list shall be developed, maintained and centrally controlled by Commandant (G-SIA). Appendix B is the approved list of Coast Guard Class Words for inclusion in the CG DADS as of the date of this publication.

2.1.3 Test for Class Words: To assist in identifying Class Words, the following question can be asked:

What type of information is being maintained about the primary object in the data element name? For example:

VESSEL NAME

The primary object described is a "VESSEL". The word "NAME" describes the type of information being maintained.

To help determine the Class Word in this example, ask the question: What type of information is being maintained about the vessel? For this example, the answer is: The name of a vessel.

VESSEL NAME
(Prime/Class Word)

2.2 Prime Words

2.2.1 Definition of Prime Words: The Prime Word identifies the object or element being defined. Typically, these objects represent a person, place, thing or event about which an organization wishes to maintain information. Prime Words may act as primary search identifiers when querying a database system and provide a basic list of keywords for developing a general-to-specific classification scheme based on business usages.

2.2.2 Prime Word Examples: Since the volume of required Prime Words is likely to number in the thousands, a preliminary Prime Word list must be developed as a restricted vocabulary for coordination and use across the organization. This list can be developed from a review/refinement of terms appearing in the names of data elements defined in existing data systems. The premise is

that our data requirements are relatively stable; most of the Prime Words already exist in the names of data elements in data systems already developed. Commandant (G-SIA) shall request a preliminary list of Prime Words from the Coast Guard's network of IRM personnel for entry into the CG DADS. This preliminary list shall be continually reviewed and updated.

The following are examples of Prime Words:

AGENCY	AIRPORT	BIRTH	CARGO
COMMAND	COUNTRY	CUTTER	DELIVERY
DESTINATION	EQUIPMENT	MESSAGE	MISSION
PERSON	PRODUCT	REPORT	REQUEST
TARGET	UNIT	VESSEL	WEATHER

An official Prime Word list, to be used in constructing CG data element names, shall be developed, maintained and centrally controlled by Commandant (G-SIA).

2.2.3 **Tests for Prime Words**: To aid in identifying Prime Words, the following questions can be asked:

What is being described?
To what does the Class Word (and its modifiers) apply?

For example: **ADMINISTRATIVE UNIT IDENTIFIER**

The primary object is "ADMINISTRATIVE UNIT". The word "IDENTIFIER" describes the type of information being maintained.

To help determine the Prime Word in this example, ask: CLASS WORD OF WHAT? IDENTIFIER OF WHAT?

For the above example, the answer is: IDENTIFIER OF A UNIT.

ADMINISTRATIVE UNIT IDENTIFIER
(Mod /Prime/Class)

2.3 Modifiers and Qualifiers

2.3.1 **Definitions**: A Modifier gives additional information about the Class or Prime Word. Modifiers may be adjectives or nouns. A Qualifier is a special kind of Modifier that is used with a Class Word to further describe a characteristic of the Class Word within a domain of values or type of information which can be attached to an object. Examples of qualifiers include specific units of

measure. The maximum number of Qualifiers allowed to describe the Class Word is two. The data element name shall not exceed **NINE** words.

2.3.2 Modifier Examples:

ANNUAL	BEGIN	BEGINNING	DAILY
DAMAGED	DESTROYED	END	ENDING
FIRST	HIGHEST	LAST	LEAST
LOWER	LOWEST	MAXIMUM	MINIMUM
MOST	ORDERED	QUARTERLY	REQUESTED
REQUIRED	START	THIRD	UPPER
VERIFIED			

Qualifier Examples:

FEET	INCHES	KILOMETERS
MILES	SECONDS	WEEKS

2.3.3 Tests for Modifiers: To aid in identifying Modifiers, the following questions can be asked:

What kind of (insert Class Word)?
What kind of (insert Prime Word)?

EXAMPLE 1: **CIVILIAN EMPLOYEE START DATE**

The word "CIVILIAN" describes the use or purpose of the object being described, in this case "EMPLOYEE". The word "START" further describes the Class Word "DATE."

To help determine the Modifiers, ask:

What kind of employee? The answer is: **CIVILIAN EMPLOYEE**
What kind of date? The answer is: **START DATE**
EXAMPLE 2: **CUTTER PATROL DISTANCE MILES**

Prime Word: **CUTTER** Modifier : **PATROL**
Class Word: **DISTANCE** Qualifier: **MILES**

PATROL modifies the Class Word/ DISTANCE. MILES further qualifies the Class Word.

These examples specifically address guidelines for cohesion and integrity. This construction readily adapts to the proposed approach for developing a classification scheme usable in establishing the uniqueness of a data element. By requiring the data element name to have one Prime Word, the data element is formulated to describe **only one concept** (cohesion). By restricting the data element name to one Class Word the data element will have only one use or meaning (integrity).

2.4 Construction/Reconstruction of Data Element Names:
The rules provided in this document shall be applied to construction of data element names for new data requirements. However, in some cases the rules may also be applied to reconstruction of names for existing data elements. When information sharing capability and compatibility are required, data element names must adhere to the same principles and structural rules; otherwise common data elements cannot be identified and associated. In this case, the established rules must be applied to rename existing data elements. However, the integrity and comprehension of the names must be maintained. In addition, reconstructed names from existing data elements can serve as a baseline for the development of data elements for new systems.

2.5 Enforceability: Automated edit/validity checks will help enforce data naming standards. As new data element definitions are added to the CG DADS, they must be edited for conformity to standards. Each data element name shall be subjected to a series of automated tests. These tests shall include:

- Absence of Class Word
- Absence of Prime Word
- Minimum Number of Words (2)
- Maximum Number of Words (9)

Tables of Prime Words and Class Words shall be centrally controlled by the Data Administration staff, G-SIA and made available in edit tables as a part of the automated tools for construction/reconstruction of data element names. As new standard words are approved, edit tables shall be updated.

SECTION 3: DATA ELEMENT STANDARDIZATION RULES

3.1 Data Element Definition Formulation Rules: This section describes the basic construction criteria for data element definitions. **The development of good definitions is crucial to the development of meaningful names**. The definition of a data element is based on logical, not physical, characteristics. This means that the definition shall reflect the purpose of the data element; i.e., what the data element is, rather than how, where or when it is used. **Data elements shall be independent of any particular hardware, software and organizations using the data**. Data elements defined in this manner may be used by other organizations or other applications than those for which they were originally intended. This is a crucial

requirement for supporting our strategy for the future, in which data should be shared to the greatest extent possible across organizations and applications. To support this requirement, the following rules shall be used to construct data element definitions:

1. Each data element must have a brief, precise, understandable definition.

2. The definition must have a meaning significantly different from that of any other data element.

3. Data element definitions must have only one acceptable interpretation.

4. Significant terms with different/varying connotations must have their use clearly explained in the definition.

5. Data element definitions must be functionally correct interpretations, written in commonly used language.

6. Functionally familiar terms, which do not literally state what the data element is or which may be uncommon to other potential users, must be avoided.

7. Use precise, consistent, grammatically correct language and avoid the use of local and repetitious expressions.

8. Do not develop restrictive definitions for information that could be used in other activities, organizations, functions, etc. While the originator's primary concern lies with his use of a data element, consideration must be given other users who may require the same or varying applications of the data element.

9. Should the inclusion of restrictive information in a definition be required, the restrictive information should be preceded by the abbreviation "e.g.," to cite an example using the restrictive information or "i.e.," to explain or further clarify use of restrictive information.

10. Data elements are logical rather than physical. Definitions must state "What the data element is", **NOT**: "How it is used", "Where it is used", "When it is used" and/or "How it is constructed."

11. Do not include processing/editing criteria in the definition.

12. The actual data element name must not be repeated in its own definition.

13. The Class Word must appear within the text of the definition.

3.2 Data Element Naming Rules: This section addresses rules for constructing data element names. Appendix C provides issues and examples in the use of these naming rules.

Rule 1: Data element names shall be derived from the data element's definition. Where possible, develop the definition using a single, simple sentence. The data element name must reflect the contents of the definition. Definitions must be clear, complete and comprehensible by CG personnel. Guidance on the formulation of data element definitions is provided in section 3.1 above.

Rule 2: Class Words are reserved; i.e. they shall not be used as Modifiers or Prime Words.

Rule 3: Each data element name shall have **one Class Word** taken from a pre-defined Class Word list. Class Words may be added to the standardized list upon approval by Commandant (G-SIA).

Rule 4: Each data element name shall have **one Prime Word** taken from a pre-defined Prime Word list. The Prime Word must be positioned as the first or second word in the data element name. Prime Words may be added to the standardized list upon approval by Commandant (G-SIA).

Rule 5: Each data element name shall contain a sufficient number of Modifiers and Qualifiers to fully describe it (up to four modifiers per Prime Word and one modifier plus two qualifiers per Class Word).

Rule 6. The sequence of words in a data element name are: **(1) Modifier** if required + **Prime Word + (1 to 4) Modifier(s)** if required + **Class Word + (1 or 2) Qualifier(s)** if required. **The Prime Word must precede the Class Word within in a data element name. The minimum number of words in a data element name is two (Prime Word + Class Word), and the maximum number of words is nine**.

Rule 7: Data element names shall not contain articles, conjunctions or prepositions. Plurals of Class Words/Prime Words are not permitted.

Rule 8: Abbreviations/acronyms shall not be permitted unless: (1) a data element name exceeds the maximum length of 250 characters; or (2) the abbreviation/acronym is so commonly used in place of the whole expression that the expression is better known by its abbreviated form.

Rule 9: Only alphanumeric characters are permitted to create the words in a data element name, with blanks used to separate the words in the name. Hyphens are permitted only when used to create new words necessary to represent the exact meaning of a data element (see examples in Appendix C). Exceptions: Numbers may not be used as the first character of a word in a data element name.

Rule 10: Block titles, line titles and column headings contained on forms, listings or reports shall not be used as data element names unless such titles or headings fulfill the requirements of a data element name as described herein.

Rule 11: Directives, forms, report control symbols, screens and product control numbers shall not be used in data element names since they are restrictive and subject to change.

Rule 12: Data element names shall not reference a specific computer, information system or organization name.

APPENDIX A: DATA ELEMENT STANDARDIZATION LIFE CYCLE

Each standard data element will have a standardization status assigned to it in the Coast Guard Data Administration Dictionary System (DADS). These elements include: development element, candidate element, approved element, installed element, and archived element. Each status marks the progress of the element through the data standardization process. Following is a description of each status:

1. **Development Element**. The status assigned to a data requirement that has been identified and defined by the developer but has not yet been submitted to Commandant (G-SIA) for review. A development status element can be a data element or a change to an existing data element. In addition, a development status element can be a data element alias. The system developer may modify an element with this status.

2. **Candidate Element**. The status assigned to a data requirement that has been identified, defined and submitted to Commandant (G-SIA) for review. Only the Data Administration staff can modify or change the status of a Candidate element.

3. **Approved Element**. Candidate elements which have passed all reviews are upgraded to approved elements by Commandant (G-SIA). Data values based on approved elements may be used for development.

4. **Installed Element**. Once an approved element is used in a production application, it is upgraded to an installed element.

5. **Archived Element**. Installed elements become archived elements when they no longer support a data requirement. The archived element and its associated attributes will remain in the data element database for a period of time established by Commandant (G-SII).

In addition to the above, a data element that has been submitted to Commandant (G-SIA) and is rejected, or disapproved, is given a "No" status. This is the same as Development status. The element is returned to the submitter for modifications and resubmission.

APPENDIX B: OFFICIAL CLASS WORD LIST

CLASS WORD	DEFINITION
ALTITUDE	THE VERTICAL DISTANCE OF A LEVEL, A POINT, OR AN OBJECT CONSIDERED AS A POINT, MEASURED FROM MEAN SEA LEVEL. (INCLUDES ELEVATION)
AMOUNT	A MONETARY VALUE ARRIVED AT BY COUNTING, AGGREGATION, OR CALCULATION.
ANGLE	THE ROTATIONAL MEASUREMENT BETWEEN TWO LINES DIVERGING FROM A COMMON POINT.
CODE	SEQUENCE OF CHARACTERS (COMBINATIONS OF ONE OR MORE NUMBERS, LETTERS, SPECIAL CHARACTERS) WHICH IS SUBSTITITED FOR A SPECIFIC MEANING. [MUST HAVE A SPECIFIC DOMAIN.]
COORDINATE	A STRING OF ALPHA-NUMERIC CHARACTERS THAT DESIGNATES THE LOCATION OF A LINE OR PLANE. (INCLUDES LONGITUDE, LATITUDE)

COUNT	A NON-MONETARY NUMERIC VALUE ARRIVED AT BY COUNTING, AGGREGATION, OR CALCULATION.
DATE	THE NOTATION OF A SPECIFIED PERIOD OF TIME.
DIMENSION	A MEASURED LINEAR DISTANCE [ONE-DIMENSIONAL]. (INCLUDES DEPTH, DIAMETER, HEIGHT, LENGTH, RADIUS, WIDTH)
DISTANCE	THE EXTENT OF ADVANCE AWAY OR ALONG FROM A POINT CONSIDERED PRIMARY OR ORIGINAL.
DURATION	AN INTERVAL OF TIME.
FACTOR	A QUANTITY BY WHICH A GIVEN QUANTITY IS MULTIPLIED OR DIVIDED IN ORDER TO INDICATE A DIFFERENCE IN MEASUREMENT.

CLASS WORD	DEFINITION
HEADING	THE COURSE OR DIRECTION (NORMALLY IN WHICH SOMETHING POINTS OR MOVES.)
HUMIDITY	A MEASURE OF THE QUANTITY OF WATER VAPOR IN THE AIR.
IDENTIFIER	A SEQUENCE OF CHARACTERS (COMBINATION OF ONE OR MORE NUMBERS, LETTERS, SPECIAL CHARACTERS) WHICH DEFINE A SPECIFIC OBJECT. [MUST HAVE A GENERAL DOMAIN.]
INDICATOR	A CONDITION CAPABLE OF HAVING ONE OF TWO VALUES.
MASS	THE MEASURE OF INERTIA OF A BODY.
NAME	A DESIGNATION FOR AN OBJECT EXPRESSED IN A WORD OR WORDS.
NUMBER	A NUMERAL OR COMBINATION OF NUMERALS OR OTHER SYMBOLS USED TO IDENTIFY OR DESIGNATE.
QUANTITY	A NON-MONETARY NUMERIC VALUE. (INCLUDES AVERAGE, BALANCE, DEVIATION, INDEX, LEVEL, MEAN, MEDIAN, MODE).
RATE	A QUANTITY, AMOUNT, OR DEGREE OF SOMETHING MEASURED PER UNIT OF SOMETHING ELSE. (INCLUDES ACCELERATION, DENSITY, POWER, PRESSURE, TENSION, TORQUE, VELOCITY, VISCOSITY)
SIZE	A REPRESENTATION OF THE PHYSICAL DIMENSIONS, PROPORTIONS, OR EXTENT OF AN OBJECT. (INCLUDES SCALE)
SPEED	THE MAGNITUDE OF THE DISTANCE PER UNIT TIME THAT AN OBJECT TRAVELS.

CLASS WORD	DEFINITION
TEMPERATURE	A MEASURE OF THE HOTNESS OR COLDNESS OF SOMETHING.
TEXT	AN UNFORMATTED CHARACTER STRING, GENERALLY IN THE FORM OF WORDS.
TIME	A NUMBER REPRESENTING A SPECIFIC POINT ON THE IRREVERSIBLE SUCCESSION OF REAL LIFE EVENTS.
VOLUME	THE AMOUNT OF SPACE OCCUPIED BY A THREE DIMENSIONAL FIGURE.
WEIGHT	THE FORCE WITH WHICH AN OBJECT IS ATTRACTED TOWARD THE EARTH BY GRAVITY.

APPENDIX C: ISSUES AND EXAMPLES CONCERNING THE NAMING RULES

The next several pages provide examples and issues pertinent to the data element naming rules defined in Section III. The examples are actual data elements found in CG data inventories. In each example, the Prime Word, Class Word and any Modifiers are identified. Actual examples are presented to focus attention on current, existing problems.

Rule 1: Data element names shall be derived from the data element's definition. Where possible, develop the definition using clear, simple sentences. The data element name must reflect the contents of the definition. Definitions must be clear, complete and comprehensible by CG personnel.

Comment: Definitions shall not be derived from data element names. A data element name can be identified more accurately after its definition is developed. By restricting definitions to clear, simple sentences, those developing data elements are encouraged to define them in a cohesive manner.

Example 1: ORIGINATOR

Definition: Any combination of characters which is the unique identifier of the originator of a message.

Reconstructed DE: **MESSAGE ORIGINATOR IDENTIFIER**

(Mod /Prime /Class)

Example 2: UNIT TYPE

Definition: Data field that is the identifier of the classification of a unit.

Reconstructed DE: **UNIT CLASSIFICATION IDENTIFIER**
(Mod /Prime /Class)

Rule 2: Class Words are reserved; i.e. they shall not be used as Modifiers or Prime Words.

Comment: A sufficient number of words and the avoidance of confusing abbreviations contribute to better naming standards.

Example 1: DATE OF BIRTH

Definition: Date of birth for an Employee
(MMDDYY)

Reconstructed DE: **EMPLOYEE BIRTH DATE**
(Prime /Mod /Class)

Rule 3: Each data element name shall have **one Class Word** taken from the pre-defined Class Word list.

Comment: Missing Class Word - By restricting the data element name to one Class Word, the data element is formulated to describe only one use. If a data element name does not contain a Class Word, derive the Class Word from the definition.

Example 1: TASK AUTHORIZATION REQUIRED DESCRIPTION
(Mod /Prime /Mod /Mod)

Definition: Text field which describes the required task authorization.

Reconstructed DE:
AUTHORIZATION REQUIRED DESCRIPTION TEXT
(Prime /Mod /Mod /Class)

Example 2: MEMBER OCCUPATION
(Prime /Mod)

Definition: The military speciality and billet, or civilian occupation and position, of a military or civilian member of the Coast Guard.

Reconstructed DE: **MEMBER OCCUPATION NAME**
(Prime /Mod /Class)

Comment: Use of multiple Class Words creating multiple ideas - Data element names containing two or more separate ideas must be broken down into individual data elements containing one idea and one Class Word.

Example 4: EVENT INITIATION DATE AND TIME
(Mod /Class /Class)

Definition: The date and time at which an event is initiated.

Reconstructed DE: **EVENT INITIATION DATE**
(Prime/Mod /Class)

Reconstructed DE: **EVENT INITIATION TIME**
(Prime/Mod /Class)

Rule 4: Each data element name shall have **ONE Prime Word** selected from a pre-defined Prime Word list. The Prime Word must be positioned as the first or second word in the data element name.

Comment: By requiring a data element to have one Prime Word, it is formulated to explicitly describe only one concept. Missing Prime Words are derived from the definition and properly inserted into the data element name.

Example 1: CITY NAME
(Mod /Class)

Definition: The city name of a CG's unit geographic location as defined by the U.S. Postal Service.

This definition is sufficient to create a generic standard data element name if we make "CITY" the Prime Word. There would, however, be many aliases because many applications/systems use this kind of information. It is better to construct the data element so that it retains enough generality to be used across applications, but specific enough as to be practical for effective corporate data administration.

Reconstructed DE: **UNIT LOCATION CITY NAME**
(Mod/Prime /Mod /Class)

Example 2: DISTANCE MILES
(Class /Qual)

Definition: A distance measured in miles.

The question "distance of what?" has not been answered by this definition. The Prime Word is missing. To create a standard data element name from this definition would require additional information. There is, therefore, no way to reconstruct this example name with the definition as is. An example of a good definition would be:

Flight distance of an aircraft measured in miles
Reconstructed DE: **AIRCRAFT FLIGHT DISTANCE MILES**
(Mod /Prime /Class /Qual)

Comment: Aggregate data elements - The following examples qualify as aggregate data element groups. Data elements contained in an aggregate group are usually related by a common Prime Word. To maintain the relationships among data element names and to allow data element names to maintain their validity, the core Prime Word should appear in each data element name within the group. Data elements contained in a database table can very often be represented in this way. Each of the data element names must be able to stand on their own with definitions that accurately define that data element.

Example 1: EQUIPMENT CATEGORY NAME AND MODEL
(Prime /Class /Class /Prime)
Note that the following data elements are contained in the same aggregate group related to the common Prime Word EQUIPMENT:

Definition: A description of equipment.

Reconstructed DE: **EQUIPMENT CATEGORY IDENTIFIER**
(Prime /Mod /Class)
Reconstructed DE: **EQUIPMENT NAME**
(Prime /Class)
Reconstructed DE: **EQUIPMENT MODEL NUMBER**
(Prime /Mod /Class)

Rule 5: Each data element name shall contain a sufficient number of Modifiers and Qualifiers to fully describe it (up to four modifiers per Prime Word and one modifier plus two qualifiers per Class Word).

Example 1: DELIVERY DATE DESIRED
(PRIME /Class/Mod)

Definition: Desired delivery date of product requested.

Reconstructed DE: **PRODUCT DELIVERY DESIRED DATE**
(Mod /Prime /Mod /Class)

Reconstructed DE: **PRODUCT DESIRED DELIVERY DATE**
(Prime /Mod /Mod /Class)

Both of these data elements fit the definition and thus either one could be used. A Prime Word can be an event about which an organization wishes to maintain information. Therefore, "DELIVERY" can be designated as the Prime Word. Depending on the organization, either "PRODUCT" or "DELIVERY" could be the appropriate Prime Word. However, both words are necessary in the data element name to establish the correct representation of the data element. "Delivery" designates: "Class Word of what?"; "What is being described?" and yet "product" is the object that is being affected by the delivery. This example highlights the difficulty of restructuring data elements in the absence of an organization-wide data enterprise model.

Rule 6. The sequence of words in a data element name are: **(1) Modifier** if required + **Prime Word + (1 to 4) Modifier(s)** if required + **Class Word + (1 or 2) Qualifier(s**) if required. **The Prime Word must preceed the Class Word within in a data element name. The minimum number of words in a data element name is two (Prime Word + Class Word), and the maximum number of words is nine**.

Comment: Arranging the words in a data element name according to this structure improves the repeatability of data element names (i.e., the probability that two people will arrive at the same name to describe identical concepts).

Example 1: DATE OF BIRTH OF CIVILIAN EMPLOYEE
(Class /Mod /Mod /Prime)

Definition: Date of birth of a civilian employee

Reconstructed DE: **CIVILIAN EMPLOYEE BIRTH DATE**
(Mod /Prime /Mod /Class)

Example 2: WEATHER, STATE OF
(Prime /Class)

Definition: Text description of the atmospheric conditions with respect to cloudiness, precipitation, or other weather phenomena.

Reconstructed DE:

WEATHER CONDITION DESCRIPTION TEXT
(/Mod /Prime /Mod /Class)

Example 3: DECIMAL WIDTH OF SEMI-MINOR AXIS
(Mod /Class /Mod /Prime)

Definition: The decimal width dimension of a semi-minor axis of an ellipse.

Reconstructed DE:
SEMI-MINOR AXIS ELLIPSE WIDTH DIMENSION DECIMAL
(Mod /Prime/Mod /Mod /Class /Qual)

Example 5: CAUSE OF LOSS OF VESSEL
(Prime)

Definition: Text which describes the approximate cause to which the loss of specific commercial vessel is attributable.

Reconstructed DE:
REASON VESSEL LOSS DESCRIPTION TEXT
(Prime /Mod /Mod /Mod /Class)

Rule 7: Data element names shall not contain articles, conjunctions or prepositions. Plurals of Class Words/Prime Words are not permitted.

Comment: Removing plurals from data element names encourages the designer to think in terms of simple concepts and increases the probability that two people will develop the same name to describe an identical concept.

Example 1: PROJECT FUNDS CODE
(Mod /Prime /Class)

Definition: An accounting code representing project funding.

Restructured DE: **PROJECT FUND CODE**
(Mod /Prime /Class)

Example 2: COOPERATING FORCES NAME
(Modifier /Prime/Class)

Definition: Name of force which is friendly.

Restructured DE: **COOPERATING FORCE NAME**
(Mod /Prime/Class)

Example 3: ELECTRONIC COUNTERMEASURES TYPE

(Mod /Prime /Mod)

Definition: Type of electronic countermeasures (ECM) encountered or employed.

Restructured DE:
COUNTERMEASURE ELECTRONIC TYPE IDENTIFIER
(Prime /Mod /Mod /Class)

Comment: Eliminating articles, conjunctions and prepositions from data element names results in more simplified naming. It also encourages designers to avoid coupling two or more concepts in a single data element and increases the probability that two people will develop the same name for identical elements.

Example 1: DATE OF ACQUISITION
(Class /Prime)
Definition: Date of system's acquisition
Reconstructed DE: **SYSTEM ACQUISITION DATE JULIAN**
(Prime /Mod /Class/Qual)

Example 2: COUNTRY OF THE WORLD
(Prime /Mod)

Definition: The independent first-level geographic political areas and their dependencies; areas of quasi-independence; and areas with special sovereignty associations; unrecognized, but sovereign political regimes; and administrative divisions without sovereignty and outlying areas of the United States, including islands in dispute.
Reconstructed DE: **COUNTRY NAME**
(Prime /Class)

Example 3: AS OF DATE AND TIME
(Class /Class)

Definition: The reported as of date and time of an event.

Reconstructed DE: **EVENT REPORT DATE**
(Prime/Mod /Class)

Reconstructed DE: **EVENT REPORT TIME**
(Prime/Mod /Class)

Rule 8: Abbreviations and acronyms shall not be permitted unless: (1) a data element name exceeds the maximum length of 250 characters; or (2) the

abbreviation/acronym is so commonly used in place of the whole expression that the expression is better known by its abbreviated form.

Example 1: OB EQUIPMENT NAME
(Prime /Class)

Definition: Name/Nomenclature of equipment observed

Reconstructed DE: **OBSERVED EQUIPMENT NAME**
(Mod /Prime /Class)

Example 2: EMPLOYEE SSN NUMBER
(Prime /Mod /Class)

Definition: The social security number used to identify an employee.

Although "SSN" is commonly used to represent "Social Security Number", the expression itself is also used frequently. In this case, the entire expression would be the best choice for the standard name.

Reconstructed DE: **EMPLOYEE SOCIAL SECURITY NUMBER**
(Prime /Mod /Mod /Class)

Example 2: UNIT OPFAC NUMBER
(Prime /Mod /Class)

Definition: The operating facility number identifier of a Coast Guard unit. This standard name is correct as is because "OPFAC" is used much more frequently than the expression within the Coast Guard.

Rule 9: Only alphanumeric characters are permitted to create the words in a data element name, with blanks used to separate the words in the name. Hyphens are permitted only when used to create new words necessary to represent the exact meaning of a data element. Exceptions: Numbers may not be used as the first character of a word in a data element name.

Comment: By eliminating special characters, data designers are encouraged to describe data element names in terms of what the data is and not how it is stored or used. This rule improves the probability that different individuals will develop the same name for identical data elements.

Example 1:
QUALITY DEFICIENCY REPORT CONTROL NUMBER

(Prime)

Definition: A number that uniquely identifies the Quality Deficiency Report.

Reconstructed DE:
QUALITY-DEFICIENCY-REPORT CONTROL NUMBER
(Prime /Mod /Class)

Comment: Hyphens are permitted to create new words when deemed necessary. In this case, "QUALITY DEFICIENCY REPORT" is an organizational entity, and it is more meaningful to make the entire expression the Prime Word instead of the word "DEFICIENCY" or "REPORT".

Example 2: LENGTH DIMENSION OF VESSEL IN FEET
(Class /PRIME)

Definition: Length dimension of a vessel, measured in feet.

Reconstructed DE: **VESSEL LENGTH DIMENSION FEET**
(Prime /Mod /Class /Qual)

Example 3: INTERFACE UNIT, NAME OF SHIP
(Class /Prime)

Definition: The name of a ship interfacing on a data link.

Reconstructed DE: **DATA-LINK SHIP NAME**
(Mod /Mod /Prime/Class)

Comment: The words "DATA" and "LINK" are hyphenated to form one word here in order to conform to the rule that the Prime Word be the first or second word in the logical name. To place the Prime Word "SHIP" as the first word in this data element name would be confusing ("SHIP DATA LINK NAME" would appear to mean "name of a data link", which conveys the wrong meaning of the data element).

APPENDIX D: GLOSSARY

APPROVED DATA ELEMENT - Candidate data elements which have been approved for development use.

ATTRIBUTE - A characteristic, property, or description of a data element.

ARCHIVED DATA ELEMENT - Installed data elements which no longer support a data requirement.

BASELINE CONFIGURATION - A configuration that consists of an inventory of information resources (hardware, software, or data, or any combination thereof) currently operational within the organization.

CANDIDATE DATA ELEMENT - The status assigned to a new or change to an existing data element that has been defined and submitted to proponents for approval.

CLASS WORD - The class word is used to specify the type of information maintained about the Prime Word of the data element name. The Class Word identifies/describes the use or purpose of the data element.

CONCEPTUAL SCHEMA - A model of the "real world" operations and applications. In the Coast Guard, a description of the data required to support its "business functions".

DATA - A noncontextual representation of facts, concepts,, and instructions in a defined format and structure which permits processing by people or machines to derive information. Representations of people, places, things, concepts, events, or activities in a defined format and structure from which information may be derived.

DATA ADMINISTRATION FUNCTION - The DA function in the Coast Guard (G-TTC-3) has the overall responsibility for the CG's data resources and for the administration, control, and coordination of all data-related activities. DA has the responsibility for planning and defining the conceptual framework for the overall database environment.

DATA BASE - A repository for stored data that is shared and integrated. A data collection so organized for computer processing so as to optimize storage and improve the independence of the stored data structure from the processing programs.

DATA BASE ADMINISTRATOR - An individual or organizational unit generally responsible for all physical activities relating to maintenance, control, and modification of particular data bases.

DATA BASE MANAGEMENT SYSTEM - The data processing system that provides the means to store, organize, and access the information in a data base.

DATA DICTIONARY - A logically centralized repository of information about data (metadata) such as meaning, relationships to other data, origin, usage and format.

DATA ELEMENT - The most basic unit of data that can be identified and described. A data element consists of a data element name and attributes describing what it is, its representation and relationships to other entities.

DATA ELEMENT ALIAS - A data element alias identifies a data element in use in a specific information system at a specific location. It will be used to bridge current non standard names being used in fielded information systems.

DATA ELEMENT COHESION - A cohesive data element refers to a single concept. It has one, and only one, purpose.

DATA ELEMENT STANDARDIZATION - (1) The concept that the characteristics of each shared data element (i.e., name definition, ownership, usage, format, etc.) are defined uniquely and accepted by all data users across an organization. (2) The process of documenting the uniform identification, definition and representation of data in accordance with established rules.

DATA ENTITY - A class of objects of concern to an organization, about which information is maintained, usually on a computerized system. These objects are units of data or aggregates of data, for example, element, file, record, and report.

DATA INDEPENDENCE - The ability of various users (or applications) to have different views of the same data. Specifically, the way data is logically defined and utilized does not depend on a specific application, storage structure, or access strategy.

DATA INTEGRITY - The ability to preserve the completeness, currency, and accuracy of the data without unintentional changes; the ability to produce results that are correct to a predefined level and to maintain data availability; the concept that the data element will be defined and used for one, and only one, purpose.

DATA MODEL - A logical representation of a collection of data elements and the relationships among these elements. A data model can be used to represent data usage throughout an organization (enterprise model) or can represent a single data base structure (logical data base model).

DATA REPOSITORY - A knowledge bane in the form of a data dictionary that integrates information about an enterprise's business data and applications planned, current, and in development.

DATA ADMINISTRATION DICTIONARY SYSTEM (DADS) - A key computer software tool for the management of CG data and information resources. It will provide facilities for recording, storing, and processing descriptions of the Coast Guard's significant data and data processing resources.

DATA RESOURCES - A critical concept which implies that data has the same characteristics as other mmmore familiar resources, such as manpower, real property, and machinery. Like the better known resources, data resources have such characteristics as cost and value. This concept is critical to the study of data dictionary/directory systems as an enterprise tool.

DATA SHARING - The ability of a number of users to access common data. This ability helps in reducing unwanted duplication, inconsistencies and storage requirements, and speeding up processing.

DATA STEWARD - The Coast Guard office that is responsible for a particular corporate data entity or data element. This responsibility includes determining the value of the data entity and data element attributes of the data elementability of a number of users to access common data. This ability helps in reducing unwanted duplication, inconsistencies and storage requirements, and speeding up processing.

DISTRIBUTED DATA BASES - The placement of a data base, or portions of a data base, in a network environment. The logical integration of an enterprise's related data bases, which are physically stored in a network of geographically dispersed computers.

INFORMATION - A representation of facts, concepts, and assumptions derived from data foor interpretation by people or machines within the context of a subject area of study or interest.

INFORMATION REPOSITORY - The same as data repository.

KEYWORD - A word designated as being of importance for searching, accessing, or retrieving.

LOGICAL DATA ELEMENT NAME - Refers to a representation of a data element that does not depend upon physical storage or computer system characteristics.

METADATA - Information about data, that is, the description of the data resources, its characteristics, location, usage, and so on. Metadata is used to identify, describe, and define user data and their attributes.

METADATA DATABASE - A repository of stored metadata that is shared and integrated. The contents of an application's data dictionary is metadata because it contains descriptive: information on data element attributes for the data used in that application. The Coast Guard Data Administration Dictionary System (DADS) is a metadata database containing metadata about Coast Guard cross-functional data elements.

OBJECT - A person, place, thing, concept, event or activity about which an organization wishes to keep information.

PHYSICAL DATA ELEMENT NAME - A representation of the data element as it is stored or physically implemented. A field name in a data file or database is a physical data element name.

PRIME WORD - A word used in a data element name which identifies the object being described by a data element. The Prime Word represents the data groupings to which the data element belongs.